

#	Ref	Search Query	DBs	Default Operator	Plurals	Time Stamp
S1	72	(database or (data adj warehouse))	US-PGPB;	AND	OFF	2005/12/01 10:57
S2	1922	(database or (data adj warehouse))	US-PGPB;	AND	OFF	2005/11/29 09:47
S3	60	(database or (data adj warehouse))	US-PGPB;	AND	OFF	2005/11/29 09:49
S4	24	(database or (data adj warehouse))	US-PGPB;	AND	OFF	2005/11/29 09:49
S5	2	(“6961721”).Pn.	US-PGPB;	OR	OFF	2005/12/01 10:57
S6	2	(“7051277”).Pn.	US-PGPB;	OR	OFF	2007/04/26 09:18

EST Search History


[Web](#) [Images](#) [Video](#) [News](#) [Maps](#) [more »](#)

[Advanced Scholar Search](#)
[Scholar Preferences](#)
[Scholar Help](#)

Scholar All articles [Recent articles](#) Results 1 - 10 of about 29 for **"token count" weighting document similarity**. (0.18 sec)

[All Results](#)

[L Gravano](#)

[M Rorvig](#)

[P Ipeirotis](#)

[N Koudas](#)

[D Srivastava](#)

[Text joins in an RDBMS for web data integration - group of 19 »](#)

L Gravano, PG Ipeirotis, N Koudas, D Srivastava - Proceedings of the twelfth international conference on World ..., 2003 - portal.acm.org
 ... retrieval field to define the textual **similarity** between t ... of all terms and v t is a **document weight** vector ... of the well-established tf.idf **weighting** scheme from ...
 Cited by 34 - [Related Articles](#) - [Web Search](#)

[Automated assistant for organizing electronic documents - group of 3 »](#)

JO Kephart, RB Segal, SR White - US Patent 7,051,277, 2006 - Google Patents
 ... describe several **weighting** schemes and their implementations. A document is classified by computing the **similarity** between the **document** token frequencies and ...
[Related Articles](#) - [Web Search](#)

[Method for automatically producing optimal summaries of linear media - group of 3 »](#)

JT Foote, J Boreczky - US Patent 7,068,723, 2006 - Google Patents
 ... Vectors from Tokenized Text Windows **Token Count** Windows Potato (1 ... similaritymeasurementuses Term- Frequency/Inverse Document Frequency (TF/IDF) **weight** - ...
[Related Articles](#) - [Web Search](#)

[Images of similarity: A visual exploration of optimal **similarity** metrics and scaling properties of ... - group of 7 »](#)

M Rorvig - Journal of the American Society for Information Science, 1999 - doi.wiley.com
 ... two **documents** are reduced to the **token count** for corresponding ... An IDF **weight** might alter the shape of the ... 2. TREC topic-document dataset 1 by five **similarity** ...
 Cited by 31 - [Related Articles](#) - [Web Search](#) - BL Direct

[A review of Web link analysis](#)

NG Freier - staff.washington.edu
 ... measurement of index term (or **token count**) distributions whether ... is the measure of hub characteristics for the **document**. The other **weight** is the measure of the ...
[Related Articles](#) - [View as HTML](#) - [Web Search](#)

[Parallel corpora for medium density languages - group of 2 »](#)

D Varga, P Halacsy, A Kornai, MC Inc, V Nagy, L ... - Proceedings of Recent Advances in Natural Language ..., 2005 - kornai.com
 ... of raw text consisting of assumed parallel **documents**. ... sentences, normalized with the larger **token count** of the ... The relative **weight** of the two components was ...
 Cited by 2 - [Related Articles](#) - [View as HTML](#) - [Web Search](#)

[Adapting information extraction knowledge for unseen web sites - group of 3 »](#)

TL Wong, W Lam - Proceedings of the 2002 IEEE International Conference on ..., 2002 - doi.ieeecomputersociety.org
 ... classes appeared in the following fragment of a HTML document: No wSell : \$ < B ... are: 1. length: number of characters in the content 2. **token count**: number of ...
 Cited by 2 - [Related Articles](#) - [Web Search](#)

[SMS Short Form Identification and Codec - group of 2 »](#)

LM Fung - wing.comp.nus.edu.sg
 ... 19 4.4 Lexical **Similarity** Measurement ... available by **weighting** the various qualities

of x. Page 15. 9 ... Count(b) refers to the number of occurrence of bigram b. and ...
[Related Articles](#) - [View as HTML](#) - [Web Search](#)

[Experimenting to Produce a Software Tool for Authorship Attribution - group of 2 »](#)
M Cook - Undergraduate Dissertation, 2003 - dcs.shef.ac.uk
... 5.1.1: Results broken down by **document** length ... showed that there was a close **similarity** between all ... The **weighted** cusum technique, developed by Hilton and Holmes ...
Cited by 2 - [Related Articles](#) - [View as HTML](#) - [Web Search](#)

[Constraint-based Ontology Induction from Online Customer Reviews](#)
T Lee - Group Decision and Negotiation, 2007 - Springer
... the hyponym relationship), mutual exclusivity (disjoint subclasses), and **similarity** ...
vector space model for clustering **documents** (Salton and ... i) as a **weighting** ...
[Related Articles](#) - [Web Search](#)

Google ►

Result Page: 1 2 3 [Next](#)

[Google Home](#) - [About Google](#) - [About Google Scholar](#)

©2007 Google